A THERMAL ROBUST SEMICONDUCTOR DEVICE USING HFN AS METAL GATE ELECTRODE AND THE MANUFACTURING PROCESS THEREOF

BACKGROUND OF THE INVENTION

(1) FIELD OF THE INVENTION

The present invention relates in general to semiconductor devices and semiconductor manufacturing fabrication processes. More particularly, this invention presents an improved thermally robust semiconductor device having a metal electrode for both low power and high-performance sub-100 nm CMOS technologies.

(2) DESCRIPTION OF THE PRIOR ART

The whole of the semiconductor industry advancement is centered largely on the development of the device and processing techniques for its Complimentary Metal-Oxide-Semiconductor (CMOS) Field Effect Transistors (FET). In the very early days of MOSFET when aluminum was used as the metal gate, it only appears for a short period of time as the aluminum has a poor adhesion to the Silicon or Silicon di-oxide (SiO2) and high gate leakage so it was quickly replaced by the poly-silicon with heavily doped N+ dopant. Poly Silicon as a metal gate, or in short poly gate has dominated CMOS technology for more than two decades. It has the advantages of good silicon adhesion, ease of processing, and no metal diffusion or penetration problems. Furthermore, the poly gate electrode can be readily scaled down without major impact to the CMOS processing. In high performance CMOS technology, when the gate size is scaled down to 0.15 and 0.13um, dual doped gate electrodes ([p+ dopant for p-channel and n+ dopant for the n-channel) have been used to enhance its channel into surface mode. When CMOS devices are scaled further down to the sub-100nm region, the gate oxide has shrunk to less than 5 nm, and the depletion layer formed in the polysilicon gate in inversion bias becomes a significant fraction of the gate capacitance and degrades the device performance. The use of a metal gate in these CMOS devices can alleviate this problem caused by poly silicon gate associated depletion effects and dopant penetration effects. See "International Technology

Roadmap for Semiconductors", Semiconductor Industry Association, San Jose, CA, 2001 (ITRS-2001)

U.S. Patent Application serial number 2003/0197230 suggests the use of two different metals with appropriate work functions: a first metal with a first work function for the PFET area and a second metal with a second work function for the NFET area. However, this approach adds significant cost and complexity to the process. Alternatively, the same metal can be used for the gate of both the PFET area and NFET area with a mid-gap work function. While refractory metal nitrides such as TaN and TiN have been extensively investigated as the potential solutions to replace poly-Si, these materials show limited thermal stability and thus are incompatible with conventional CMOS processes (with thermal processing for activating the source and drain regions). See "Physical and electrical properties of metal gate electrodes on HfO2 gate dielectrics," by J.K. Schaeffer et. al ., Journal of Vacuum Science and Technology Vol. 21(1), Jan/Feb 2003, p.11-17 and "Thermal Stability of PVD TiN Gate and Its Impacts on Characteristic of CMOS Transistors," by M. Wang et. al. 6th International Symposium on Plasma Process Induced Damage, May 14-15 Monterey CA USA, 2001, p.36-39

The thermal stability of HfN is superior to TiN and TaN, due to its negatively larger heat of formation compared to that of TiN and TaN (HfN:–88.2, TiN:–80.4, TaN:–60.3; kcal/mol) See "Properties and microelectronic applications of thin films of refractory metal nitrides", by M. Wittmer, Journal of Vacuum Science Technology A, vol. 3, pp. 1797-1803, 1985. An attempt making use of HfNx as the metal gate has been reported. Heuss et al. in his abstract (see Heuss at al. "Thermal stability of Hafnium and Hafnium Nitride (HfNx) Gate Electrodes on Silicon Dioxide", Materials Research Society Proceedings, April 2000) discussed the use of HfNx as the metal gate material but do not recommend the application of HfN as the gate electrode. US. Patent 6,225,168 to Gardner, et al. shows a metal gate electrode and a titanium or tantalum nitride as gate dielectric barrier layer and the processes for fabricating such devices. Figure 1 shows a bulk CMOS with metal gate stack structure of Gardner et al. wherein TiN is the barrier metal and TaN is the capping layer.

US Patent 6,383,879 to Kizilyalli , et al. presents a method to form dual metal gates for the different work function for NMOS and PMOS transistors.

US patent 6,511,911 to Besser , et al. gives a metal gate stack structure comprised of Tungsten, tantalum, TiN and etch stopper which is used for the deep submicron CMOS process. Figure 2 depicts the gate stack structure of Besser et al. wherein tungsten is used as capping layer 18 and TiN as the barrier metal 14. There is second metal layer Ta 16 deposited in between the capping layer.18 and barrier layer 14.

US Patent 6,617,624 to Powell teaches a metal gate stack include a doped polysilicon, TiN and Tungsten with Nitride passivation and its formation processes

U.S. Patent 6,043,157 to Gardner et al shows a process for forming dual gates where one gate is polysilicon and the other gate is metal.

U.S. Patent 5,960,270 to Misra et al discloses a process wherein the same mid-gap work function metal is used for both n- and p-gates. U.S. Patent 6,083,836 to Rodder teaches a dummy gate process where two gates are formed. For example, one gate is polysilicon and the other is aluminum. U.S. Patent 6,051,487 to Gardner et al teaches a dummy gate process using a polysilicon or a metal gate.

SUMMARY OF THE INVENTION

A primary object of the invention is to provide a metal gate having robust thermal stability so that the equivalent oxide thickness (EOT) of the gate stack will not increase during the gate dielectric material growth.

Another object of the invention is to provide a formation process of the metal gate structure, which is compatible with the thermal annealing of the transistor's source and drain activation.

A still further object of this invention is to provide the metal gate structure having its work function tunable by varying process conditions. Yet another object of this invention is to provide a gate material that is sufficient to prevent oxygen diffusion through the gate stack during thermal processing. .

In accordance with the objects of the invention, hafnium nitride (HfN) as the gate material is presented and the formation process of such a gate is given. The semiconductor structure composition consists of at least one underlying dielectric. In common practice, it can be either the conventional SiO2 or the more recent high dielectric constant (high-K) material of HfO2, though not only limited to the two. The gate material of HfN exhibits a mid-gap work function and shows robust resistance against high temperature treatment. In particular, the equivalent oxide thickness (EOT) and gate leakage show little variation. The superior oxygen diffusion barrier property as well as the excellent thermal stability of HfN/HfO2 and HfN/SiO2 interface makes it an ideal candidate for the sub-65nm for both bulk and SOI CMOS technologies in place of the conventional poly-Si gate material. The gate structure is also ready to be implemented into the symmetrical dual gate transistor structure (SDG).

BRIEF DESCRIPTION OF THE DRAWINGS

In the accompanying drawings forming a material part of this description, there is shown:

Figure 1 is a cross-sectional representation of a bulk CMOS with metal gate stack structure in the prior art (see US patent 6,225,168 to Gardner et al.) wherein TiN as the barrier metal 22 and TaN is the capping layer.

Figure 2 depicts the gate stack structure of the prior art (see US patent 6,511,911to Besser et al.). wherein tungsten is used as capping layer 18 and TiN as the barrier metal 14. There is second metal layer Ta 16 is deposited in between the capping layer 18 and barrier layer 14.

Figure 3 is a cross-sectional representation of the preferred embodiment of the present invention as in the bulk CMOS where HfN can be used as the gate material.

Figure 4 is a cross-sectional representation of a second preferred embodiment of the invention as in a silicon-on-insulator (SOI) device.

Figure 5 shows a typical setup of physical vapor deposition (PVD) for the formation of HfN.

Figure 6 through to Figure 10 shows the process steps in forming the bulk CMOS transistor with the said HfN as metal gate electrode.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Referring now more particularly to Figure 3, there is shown a semiconductor substrate 20. This is preferably monocrystalline silicon. Isolation regions, such as shallow trench isolation (STI), not shown in the figure, are formed in the substrate as is conventional in the art to separate active regions. N-wells and P-wells, not shown, may be formed within the substrate, as is conventional.

Punchthrough and threshold voltage adjustment implantations, not shown, are made as is conventional in the art. Figure 4 illustrates an alternative where the substrate is a silicon-on-insulator substrate. All processing is the same with both substrate alternatives.

A gate dielectric layer 22 is grown or deposited over substrate to a thickness of between about 15 and 150 Angstroms. For example, the dielectric layer may be a low dielectric constant material such as silicon dioxide, nitrided silicon dioxide, silicon nitride, or their combinations. Alternatively, the dielectric layer may be a high dielectric constant gate dielectric material such as zirconium oxide, hafnium oxide, aluminum oxide, tantalum pentoxide, barium strontium titanates, and crystalline oxide.

Referring now in particular to Fig. 5, after pre-gate cleaning, the gate dielectric has been thermally grown and the substrate 40 is placed in the PVD vacuum chamber. The PVD chamber is initially set at base pressure of less than 2e-7 Torr; flowing the Nitrogen and Argon gases at constant flow rate of N2/Ar at ratio of 5 sccm/ 25 sccm, DC power set on Hf target (42) at 450V; and RF power on substrate 40 is set at 12 W. During the sputtering deposition, the gas pressure is maintained at 2 mtorr inside the chamber. This will lead to a deposition rate of approximately 8.2 nm/min.

A metal layer of HfN 26 (~50 nm) is then formed and the mid gap work function shall be at approximately 4.65eV. This HfN has the composition of Hf/N atomic ratio of 1. To obtain the excellent thermal stability of HfN, the ratio of Hafnium to nitrogen should be controlled to be less than or equal to one (equal amounts or more of nitrogen). By varying the Hf to Nitrogen ratio by way of changing the nitrogen

flow, the mid gap work function can be tuned. Alternatively the metal layer 26 may be formed by evaporation, or chemical vapor deposition (CVD).

A TaN (~100 nm) capping layer 28 is sputtered on HfN gate metal layer to achieve a low gate sheet resistance (~10 Ohm/sq.). The TaN/HfN stack is then etched using plasma dry etch method (RIE : Reactive Ion Etch) having Chlorine (Cl2) gas based chemistry. Other capping layers such as tungsten may be used.

Refer now to Figures 6 and 7. The metal and metal capping layers, 26 and 28, respectively, are deposited on the underlying dielectric and then patterned to form gate electrodes. This gate stack contains a first portion having the composition of Hafnium and nitrogen and the second portion, underlying the first portion, having the composition of Hafnium and oxygen, or silicon and oxygen or other gate dielectric whereas the second portion is in contact with the silicon substrate. A chemical mechanical polishing (CMP) process may be applied to achieve planarization. This gate stack has the capability of scale down to at least below 10Å, and could be utilized at the sub-65nm CMOS technology. The above gate stack structure after thermal treatment of 1000 °C RTA for 20 sec without using surface nitridation prior to HfO_2 deposition, can still maintain very good stability and EOT(equivalent oxide thickness) changes of less than 2 Angstroms.

Alternatively, the first metal layer 26 may be tungsten or tantalum nitride, for example. Then HfN may be deposited as described above as the capping layer 28. After photoresist 60 patterning by masking steps, the gate stack is then patterned by chlorine based dry etch to form the gate electrode having a dielectric layer 22, a first metal layer 26 on the dielectric layer, and a HfN capping layer 28 on the first metal layer.

Referring now to Figures 8 through 10, source and drain regions 30 and sidewall spacers 24 may be formed, as is conventional. The spacers 24 may comprise either silicon nitride or silicon oxide. Processing continues as is conventional in the art..

While the invention has been particularly shown and described with reference to the preferred

embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made without departing from the spirit and scope of the invention.

What is claimed is: